

Co-evolution of Content Popularity and Delivery in Mobile P2P Networks

Srinivasan Venkatramanan and Anurag Kumar
Department of Electrical Communication Engineering
Indian Institute of Science
Bangalore 560012, India
Email: vsrini,anurag@ece.iisc.ernet.in

Abstract

Mobile P2P technology provides a scalable approach to content delivery to a large number of users on their mobile devices. In this work, we study the dissemination of a *single* content (e.g., an item of news, a song or a video clip) among a population of mobile nodes. Each node in the population is either a *destination* (interested in the content) or a potential *relay* (not yet interested in the content). There is an interest evolution process by which nodes not yet interested in the content (i.e., relays) can become interested (i.e., become destinations) on learning about the popularity of the content (i.e., the number of already interested nodes). In our work, the interest in the content evolves under the *linear threshold model*. The content is copied between nodes when they make random contact. For this we employ a controlled epidemic spread model. We model the joint evolution of the copying process and the interest evolution process, and derive the joint fluid limit ordinary differential equations. We then study the selection of the parameters under the content provider's control, for the optimization of various objective functions that aim at maximizing content popularity and efficient content delivery.

I. INTRODUCTION

Peer-Peer (P2P) architectures help relieve file servers from excessive load by enabling clients to communicate among themselves and exchange content, hence providing a scalable approach to content delivery for a large audience. With the proliferation of smart phones and high speed mobile Internet, there is increasing download of media content directly into mobile devices. Mobile P2P technology aims to leverage this trend and several possible architectures [1] have been proposed. In this paper we are concerned with peer-to-peer spread of the content under the delay tolerant networking (DTN) paradigm.

While studying content distribution, it is essential to understand the evolution of demand for the content. For modeling demand evolution, we can adopt the point of view that the existing demand for an item influences others also to get interested in it. There has been considerable interest in such models in the area of viral marketing research. For example, online social networks (Facebook, Twitter, etc.) provide a platform for the users to exchange their views about a particular media content with members of their social circle and end up *influencing* each other. Content creators interested in increasing the popularity of the content use several data mining techniques to identify optimal viral marketing strategies [2] and to identify the most influential users in a social network [3]. In this paper, we are also concerned with spread of interest in the item of media content, and for this we adopt an influence spread model introduced earlier in the context of viral marketing [3].

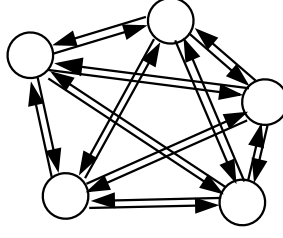
Evidently, content providers need to adopt their content dissemination strategies according to the prevailing interest in the particular content. Though there have been several works that discuss optimal strategies for content distribution in mobile ad hoc networks in the DTN setting [4],[5] and in mobile P2P [6], they do not address the issue of jointly modeling the spread of interest in the content and the content itself. In our work, we aim to model the joint evolution of demand and spread of the content in a mobile P2P setting.

A recent work [7] addresses a related problem of demand-aware distribution by employing fluid models for the viral spread of demand and aims at obtaining hybrid P2P and client-server architectures that can meet the demand. While [7] assumes that the demand spread is uncontrolled and optimizes for content delivery, in this work, we separately discuss optimizing the parameters of the demand evolution process (to increase the popularity the content), as well as the content delivery process (to efficiently meet the content demand).

Contributions: Our main contributions are:

- Developing ordinary differential equation (o.d.e.) models for the co-evolution of popularity of the content and its spread via controlled epidemic copying in a mobile P2P environment
- Using these o.d.e. approximations to provide insights into the choice of parameters for the content provider in order to optimize various system performance objectives

In Section II we introduce the combined system model and the related notations. In Section III we analyse a threshold based model for evolution of interest and derive its fluid limit for general threshold distributions. We also discuss the effect of threshold distribution on the model, and show that SIR epidemic model [8] is a special case of our model. In Section IV we derive fluid limits for evolution of spread on top of the interest evolution model. We finally provide numerical results for



$$\text{All edge weights} = \gamma, \gamma \leq \frac{1}{|\mathcal{N}|-1}$$

Fig. 1. The HILT Network Model

some practical optimizations (Section V) associated with the evolution of interest and the joint-evolution of interest and spread of the content.

II. THE SYSTEM MODEL

We consider a scenario where the users with their mobile devices, modeled as a population of mobile *nodes*, meet each other according to a random contact model. We assume that the nodes are homogeneous and we wish to model the spread of a *single* content among these nodes. As motivated in the Introduction, we aim to model the joint spread of popularity of the content and the content itself. Pertaining to the content, each node has a current state represented by two bits: the *want* bit indicates whether the node is interested in the content and the *have* bit indicates whether the node has the content. When the *want* bit of a node is 1, we call the node, a *destination* (i.e., interested in the content), else we call it a *relay* (i.e., not yet interested in the content).

There is a central server that keeps track of the *want* state of all the nodes. The total number of destinations (nodes interested in the content) is a measure of popularity of the content. In order to promote the content, this measure of popularity must be made known to all the nodes. We assume that there is a low bitrate control channel that (i) is used by the nodes to inform the central server about their interest in the content, and (ii) is used by the central server to broadcast, at regular intervals, the fraction of nodes interested in the content. Thus, for the purpose of spread of content popularity, the network is fully connected.

Relay nodes, on receiving popularity broadcasts from the central server, might get converted to destinations according to an *influence spread model*. We model this influence process (evolution of interest in the content) using the Homogeneous Influence Linear Threshold (HILT) model, a special case of the Linear Threshold model (LT) introduced in Kempe et al. [3]. For the content copying process, we model the random contacts between pairs of nodes as independent Poisson point processes (a model also used in the context of mobile P2P in [9]), with the copying (when pairs of nodes meet) being controlled probabilistically, in a manner similar to the Susceptible-Infective (SI) epidemic model [10].

A. Modeling Interest Evolution: The HILT Model

In this section we introduce the HILT model used to model the evolution of interest in the content. In the original LT model [3], nodes are part of a weighted directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. With each $i, j \in \mathcal{E}$, there is associated a weight $w_{i,j}$ which gives a measure of *influence* of node i on node j , normalized such that the total weight into any node is at most 1, i.e., $\sum_i w_{i,j} \leq 1$. The Homogeneous Influence Linear Threshold (HILT model) is a special case of the LT model where the network graph is complete and all nodes are homogeneous. Hence, we have a mesh network on the population \mathcal{N} containing $N = |\mathcal{N}|$ nodes with each edge carrying the same influence weight $\gamma_N = \frac{\Gamma}{N-1}$ and $\Gamma \leq 1$ (see Figure 1).

The evolution of interest in the content is modeled by the following influence process evolving in discrete time. Each node $j \in \mathcal{N}$ independently chooses a random threshold Θ_j from a given distribution F at the *beginning*. We begin with an initial set of destinations $\mathcal{A}(0)$. In the HILT model, the net influence of a set of destinations on any relay is γ_N times the size of the destination set. The number of destinations (currently interested nodes) is broadcast to all the nodes by the central server after each period, and a relay gets converted into a destination once the net influence exceeds its chosen threshold. We also assume the *progressive case*, i.e., conversion into destinations is an irreversible process. In other words, a relay $j \notin \mathcal{A}(k-1)$ gets influenced in step k if,

$$\gamma_N |\mathcal{A}(k-1)| \geq \Theta_j \quad (1)$$

The influence is modeled as building up cumulatively over time. The initial set $\mathcal{A}(0)$ is viewed as being *infectious* and results in some of the relay nodes “tipping” over their thresholds and getting converted to destinations in the first period. A node j that remains a relay has the level of cumulative influence on it raised to $\gamma_N |\mathcal{A}(0)| (< \Theta_j)$. The nodes in $\mathcal{A}(0)$ are now viewed as being non-infectious, and the newly infected nodes are denoted by $\mathcal{D}(1)$, with $\mathcal{A}(1) = \mathcal{A}(0) + \mathcal{D}(1)$; see Figure 2. Thus, at the end of each period the population will contain three types of nodes: $\mathcal{A}(k)$, the set of non-infectious destinations,

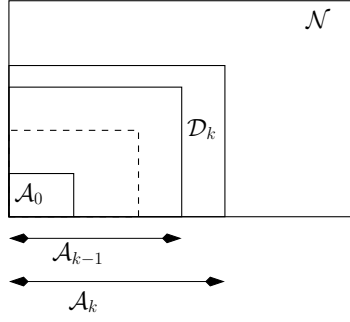


Fig. 2. Spread of Influence in the HILT Model
 $\text{want} = 0$
 $\text{have} = 0$

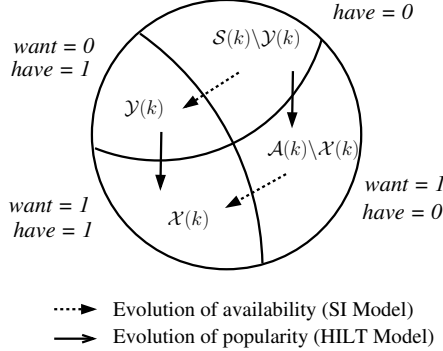


Fig. 3. Joint evolution of content popularity and spread in the HILT-SI Model.

$\mathcal{D}(k) \subseteq \mathcal{A}(k)$, the set of newly infected destinations in that period (infectious for the next period) and the set of relays denoted by $\mathcal{S}(k)$. It is clear that the activation process stops at a random step T when there are no more infectious destinations, i.e., $\mathcal{D}(T) = \emptyset$, and a *terminal set* $\mathcal{A}(T)$ is reached.

B. Modeling Content Copying: SI Model

In order to model the content delivery process, we further classify the nodes depending on whether they have the content (i.e., based on the *have* bit). Let $\mathcal{X}(k) \subseteq \mathcal{A}(k)$ denote the set of destinations that have the content, and $\mathcal{Y}(k) \subseteq \mathcal{S}(k)$ denote the set of relays that have the content.

For the evolution of $(\mathcal{X}(k), \mathcal{Y}(k))$, we model a content copying process based on the Susceptible-Infective (SI model). Between the discrete time steps, pairs of nodes meet at the points of a Poisson process with rate λ , and whenever a node that has the content meets a node that does not, content transfer takes place in a probabilistic manner. At such a meeting, we distinguish between the node that does not have the content, being a destination or a relay, by having different copy probabilities $\alpha_N = \frac{\alpha}{N}$ and $\sigma_N = \frac{\sigma}{N}$, respectively. The scaling with respect to N is to ensure that the fluid limit infection rates would be α and σ respectively [10].

C. Co-evolution of Interest and Spread: The HILT-SI Model

In the combined model, underlying the content delivery process (the SI model), the influence process (the HILT model) converts relay nodes into destinations. Thus, in our setup, the fraction of destinations is time-varying (unlike [5]). Also, the content spread is dependent on the interest evolution but not vice versa. Figure 3 shows the possible transitions between the four sets of nodes. An interesting feature of this model is the importance of copying to a relay node. As a content provider, we might be interested in delivering only to the destinations (interested in the content). But, there are two advantages of copying to a relay. First, copy to a relay promotes the further spread of the content even to destinations; this is the aspect explored in a controlled Markov process setting in [5]. Second, the relay we copy to now might later get influenced (by the HILT model) to become a destination, which is a unique feature of the HILT-SI model.

III. EVOLUTION OF INTEREST IN THE CONTENT

In this section and in Section V, we will use the convergence of a scaled discrete time Markov chain (DTMC) to a deterministic limit described by an ordinary differential equation (o.d.e.) (Kurtz [11], and also Darling [12], [13]), to develop deterministic (or, so called, fluid) approximations to the HILT model for interest evolution, and the HILT-SI model for co-evolution of interest and content spread.

Consider the HILT model on N nodes, and with edge weights $\gamma_N = \frac{\Gamma}{N-1}$. Recall that Θ_i is the influence threshold of User i , $1 \leq i \leq N$. The $\Theta_i, 1 \leq i \leq N$, are non-negative, i.i.d. random variables chosen according to the cumulative distribution

function $F(\cdot)$, a continuous distribution with density function $f(\cdot)$. We see that Γ and the threshold distribution F together define the HILT model. They can be seen as modeling the susceptance level of the population to being influenced by the popularity of the content under consideration.

Note that $\mathcal{A}(k), k \geq 1$ represents the set of destinations and $\mathcal{D}(k)$ the set of *newly added* destinations (infectious) at the end of period k with $\mathcal{D}(0) = \mathcal{A}(0)$. Define $\mathcal{B}(k) = \mathcal{A}(k-1)$, the set of destination nodes up to period $k-1$ with $\mathcal{B}(0) = \emptyset$. Thus, for $k \geq 1$, $\mathcal{B}(k) = \cup_{0 \leq i \leq k-1} \mathcal{D}_i$. We will be working with sets $\mathcal{B}(k)$ and $\mathcal{D}(k)$ to derive the fluid limits of the HILT process. Also, since the nodes are homogeneous in the HILT model, it suffices to record the sizes of the respective sets, and not on the exact membership of the sets themselves. Let $A(k)$, $D(k)$ and $B(k)$ be the sizes of the sets $\mathcal{A}(k)$, $\mathcal{D}(k)$ and $\mathcal{B}(k)$ respectively.

A. O.D.E. Model for Interest Evolution

We can show that the original HILT process $(B(k), D(k))$ is a Markov chain (see Appendix A). In order to obtain an approximating o.d.e., we work with an appropriately scaled Markov process $(B^N(t), D^N(t))$, which evolves on a time scale N times faster than that of the original system. We can visualize this process as evolving over “minislots” of duration $1/N$, whereas the original process evolves at the epochs $k = 0, 1, 2, \dots$. The minislots are indexed by $t = 0, 1, 2, \dots$. Since this new process runs on a faster time scale, we slow down its dynamics by utilizing an approach taken in [14]. In the present context, the scaling approach can be interpreted as follows. In each minislot, each infectious destination in $D^N(t)$ is permitted to influence the relays with probability $\frac{1}{N}$ and its influence is deferred with probability $1 - \frac{1}{N}$. In the former case, it contributes its influence of $\frac{\Gamma}{N-1}$ and then moves to the set $B^N(t+1)$, otherwise it stays in the $D^N(t+1)$ set. In the original process, the influence of all the newly infected destinations will be taken into account by the central server while announcing the popularity level of the content in the next time step, whereas, in the scaled process, only those infectious destinations that choose to use their influence will be considered. Define by $C^N(t) \subseteq D^N(t)$ the set of infectious destinations who use their influence at time t . Then,

$$\begin{aligned} C^N(t) &= \frac{D^N(t)}{N} + Z_b^N(t+1) \\ B^N(t+1) &= B^N(t) + C^N(t) \end{aligned}$$

and by applying Equation 1 to the relay nodes $j \in \mathcal{N} \setminus \mathcal{A}(k)$,

$$\begin{aligned} D^N(t+1) &= D^N(t) - C^N(t) \\ &+ \mathbb{E} \left[\frac{F(\gamma_N(B^N(t) + C^N(t))) - F(\gamma_N B^N(t))}{1 - F(\gamma_N B^N(t))} \right] \times \\ &\quad \left(N - B^N(t) - D^N(t) \right) + Z_d^N(t+1) \end{aligned}$$

where $Z_b^N(t+1)$ and $Z_d^N(t+1)$ are zero mean random variables conditioned on the history of the process $(B^N(0), D^N(0), \dots, B^N(k), D^N(k))$ representing the noise terms, and the expectation in the expression for $D^N(t+1)$ is with respect to $C^N(t)$. Define $\tilde{B}^N(t) = \frac{B^N(t)}{N}$ and similarly $\tilde{C}^N(t)$ and $\tilde{D}^N(t)$ to denote the fraction of users of each type (destination and relay). $(\tilde{B}^N(t), \tilde{D}^N(t))$, $t = 0, 1, 2, \dots$ is a DTMC on the state space $[0, \frac{1}{N}, \frac{2}{N}, \dots, 1] \times [0, \frac{1}{N}, \dots, 1]$ with $\tilde{B}^N(t) + \tilde{D}^N(t) \leq 1$.

Then we can state the following theorem on the convergence of the scaled process to their deterministic limits.

Theorem 1: Given the interest evolution Markov process $(\tilde{B}^N(t), \tilde{D}^N(t))$, with bounded $\dot{f}(\cdot)$ for the threshold distribution, we have for each $T > 0$ and each $\epsilon > 0$,

$$\mathbb{P} \left(\sup_{0 \leq u \leq T} \left| (\tilde{B}^N(\lfloor Nu \rfloor), \tilde{D}^N(\lfloor Nu \rfloor)) - (b(u), d(u)) \right| > \epsilon \right) \xrightarrow{N \rightarrow \infty} 0$$

where $(b(u), d(u))$ is the unique solution to the o.d.e.,

$$\begin{aligned} \dot{b} &= d \\ \dot{d} &= \frac{f(\Gamma b) \Gamma d}{1 - F(\Gamma b)} (1 - b - d) - d \end{aligned}$$

with initial conditions $(b(0) = 0, d(0) = a(0))$.

Proof: This is essentially an instance of Kurtz’s Theorem [11]; see also [12], [13]. In Appendix E we derive equivalent conditions for verifying Kurtz’s theorem. Refer Appendix B for the steps involved in deriving the fluid limit. A detailed proof verifying the necessary conditions for Kurtz’s theorem for the HILT case is provided in Appendix F-A. ■

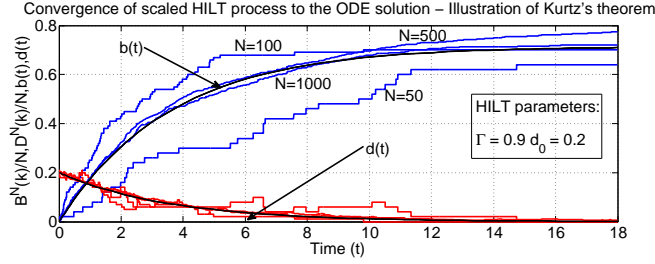


Fig. 4. Convergence of the scaled HILT Markov chain $(\tilde{B}^N(k), \tilde{D}^N(k))$ to the o.d.e. limit $(b(t), d(t))$. The o.d.e. solution $(b(t), d(t))$ plotted along with sample paths of the scaled process $(\tilde{B}^N(k), \tilde{D}^N(k))$ for $N = 50, 100, 500, 1000$.

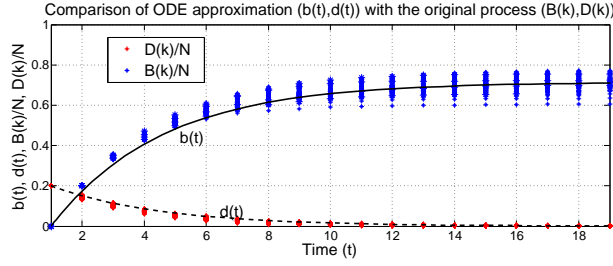


Fig. 5. Multiple sample paths of the unscaled HILT process $(B(k), D(k))$ for $N = 3000$ is plotted along with the o.d.e. solution $(b(t), d(t))$ for $\Gamma = 0.9$ and $d_0 = 0.2$

Remark: We know that the hazard function[15] for the cdf $F(x)$ is given by $h_F(x) = \frac{f(x)}{1-F(x)}$ where $f(x)$ is the corresponding probability density function. Hence the o.d.e. becomes,

$$\dot{b} = d \quad (2)$$

$$\dot{d} = h_F(\Gamma b)\Gamma d(1-b-d) - d \quad (3)$$

B. Accuracy of the O.D.E. Approximation

Consider the HILT process of interest evolution, under the special case of uniform distribution of influence thresholds, i.e., $\Theta_i \sim U[0, 1]$. The hazard function corresponding to uniform distribution is given by

$$h_F(x) = \frac{1}{1-x}$$

The corresponding o.d.e.s then become:

$$\dot{b} = d \quad (4)$$

$$\dot{d} = \frac{\Gamma d}{1-\Gamma b}(1-b-d) - d \quad (5)$$

Figure 4 shows the convergence of the scaled HILT process $(\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor))$ to the solutions of the above o.d.e. $(b(t), d(t))$ for increasing values of N (50, 100, 500, 1000). Γ and d_0 were chosen to be 0.9 and 0.2 respectively. We see that for $N = 500, 1000$ the o.d.e. approximates the scaled HILT Markov chain very well.

The original HILT process of interest evolution $(B(k), D(k))$ is then compared with the solutions of the o.d.e. $(b(t), d(t))$. The results are shown in Figure 5, where multiple sample paths of the original discrete time HILT process are superimposed on the o.d.e. solutions. We find that the o.d.e. solution approximates the original process really well, and hence permits using the fluid limit approximation for sufficiently large N .

C. Effect of the Threshold distribution

In the HILT model, while Γ is indicative of the total level of influence each individual can receive from the others, the threshold distribution $F(\cdot)$ captures the variation among the individuals' susceptance levels for getting interested the content. An empirical analysis on the effect of threshold distributions on collective behavior is available in [16]. Having established the o.d.e. limit and studied its ability to approximate the evolution of interest process, in this section we exploit the o.d.e. limit to study the effect of the threshold distribution.

1) *Uniform Distribution:* For the HILT model of interest evolution under the uniform threshold distribution, we can state the following theorem.

Theorem 2: Given the initial fraction of destinations d_0 , in an HILT network with influence weight Γ under uniform distribution, define $r = 1 - \Gamma + \Gamma d_0$. Then,

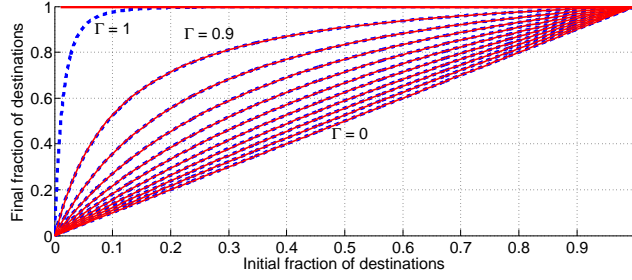


Fig. 6. Final fraction of destinations as predicted by the fluid limit b_∞ (bold lines) and by the discrete formulation $\frac{I_{\gamma_N}(k)}{N}$ (dashed lines) plotted for various values of Γ and d_0 .

- 1) The fluid limit for the evolution of interest is given by,

$$b(t) = \frac{d_0}{r} - \frac{d_0}{r} e^{-rt} \quad (6)$$

$$d(t) = d_0 e^{-rt} \quad (7)$$

- 2) The final fraction of destinations is given by

$$b_\infty = \frac{d_0}{1 - \Gamma + \Gamma d_0} \quad (8)$$

Proof: The first part of the theorem is obtained by explicitly solving the o.d.e.s for the HILT model under uniform distribution (Equations 4 and 5) for the initial conditions $b(0) = 0, d(0) = d_0$ (see Appendix G). ■

Linear Threshold model under uniform distribution has been studied in discrete setting for general networks in [3]. Since in the HILT model all nodes are homogeneous, we will be interested in the influence of a set of size k . Consider the HILT network with N nodes and influence weight γ_N . Let $I_{\gamma_N}(k)$ be the expected size of the *terminal set of destinations* $\mathcal{A}(T)$, starting with \mathcal{A}_0 of size k as the initial set of destinations. By using results from [17], we can show that,

$$I_{\gamma_N}(k) = k[1 + (N - k)\gamma_N[1 + (N - k - 1)\gamma_N[1 + \dots$$

In the expression for $I_{\gamma_N}(k)$, noting $\lim_{N \rightarrow \infty} N\gamma_N = \Gamma$ and $d_0 = \frac{k}{N}$, we can show that as $N \rightarrow \infty$, $\frac{I_{\gamma_N}(k)}{N} \rightarrow b_\infty = \frac{d_0}{r}$ (refer Appendix H). This reconfirms the fact that the fluid model is consistent with the discrete formulation. Also, while [17] allows us to compute only the final fraction of destinations, our current work provide a good approximation of the actual trajectories of the influence process.

Remarks: Figure 6 shows the behaviour of b_∞ for various of Γ and d_0 . Plotted alongside are the values of $\frac{I_{\gamma_N}(k)}{N}$ for the corresponding values of γ_N and k for $N = 3000$ and it can be seen that the two solutions match really well. It is easy to note that when $\Gamma = 0$, the HILT process does not have any effect on the number of destinations, i.e., the fraction of destinations and relays in the population remains a constant, similar to [5]. We observe from the fluid limit that, as long as $\Gamma < 1$ we cannot influence the entire population (i.e., $b_\infty < 1$) unless we start off with the entire population to be destinations (i.e., $d_0 = 1$). But if $\Gamma = 1$, then the influence of the destinations are at the maximum, and we can ultimately convert all nodes into destination, given we start with a non-zero fraction of destinations, i.e., $b_\infty = 1$ provided $d_0 > 0$.

D. Exponential Distribution

For the uniform distribution the hazard rate function is $h_F(x) = \frac{1}{1-x}$ which is monotonically increasing in x . Such a hazard rate implies that the population consists of relay nodes are more susceptible to getting interested in the content as the total number of destinations increases (i.e., total accumulated influence over the past increases). The exponential distribution with parameter Λ yields $h_F(x) = \Lambda$, a constant hazard rate function. This implies a *memoryless* property for the influence process, i.e., the relay nodes are equally likely to get influenced at a given time instant, irrespective of the net accumulated influence in the past. We can then state the following theorem.

Theorem 3: Given the initial fraction of destinations d_0 , in an HILT network with influence weight Γ under exponential distribution with parameter Λ ,

- 1) The fluid limit of the evolution of interest is given by the solution to the o.d.e.,

$$\begin{aligned} \dot{b} &= d \\ \dot{d} &= \Lambda\Gamma d(1 - b - d) - d \end{aligned}$$

- 2) The final fraction of destinations is the solution to the transcendental equation,

$$b_\infty = 1 - (1 - d_0)e^{-\Lambda\Gamma b_\infty}$$

Proof: The first part is obtained by substituting $h_F(x) = \Lambda$ in the HILT o.d.e. (Equations 2 and 3). This is equivalent to the SIR epidemic model with infection rate $\lambda\Gamma$ and recovery rate of 1 [18]. Note that $b(t)$ is then equivalent to the Recovered

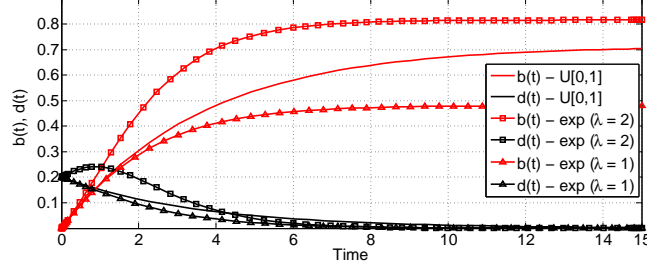


Fig. 7. Effect of threshold distribution in the HILT Model. Plotted are the interest evolution trajectories corresponding to Uniform distribution and Exponential distribution ($\lambda = 1$ and $\lambda = 2$).

set (R), and $d(t)$ is equivalent to the Infected set (I). The second part follows from classic SIR literature [8] and observing that the basic reproduction number (expected number of new infections from a single infection) is $\Lambda\Gamma$. ■

Remarks: Figure 7 shows the comparison of the interest evolution process ($b(t), d(t)$) with $\Gamma = 0.9$ and $d_0 = 0.2$, for uniform $[0, 1]$ and exponential distribution with parameter $\Lambda = 1, 2$. Observe that, when the expected values of the thresholds are the same (uniform $[0, 1]$ and exponential with $\Lambda = 2$), the exponential distribution results in a higher value of b_∞ , since there will be a larger fraction of relay nodes with low sampled threshold. Also since $\Lambda\Gamma > \frac{1}{1-d_0}$, $\dot{d}(0) > 0$ and in SIR terminology[8] a *proper outbreak* is said to have occurred. As Λ decreases, we see that the expected influence thresholds of the relay nodes increase, and hence we notice a significant drop in the interest evolution process.

IV. JOINT EVOLUTION OF INTEREST AND SPREAD OF THE CONTENT

In the previous sections we studied the evolution of interest in the content. In this section, we shall adopt similar techniques to obtain the joint evolution of interest and spread of the content (HILT-SI model) in mobile P2P setting. Recall that interest evolution is modeled by the HILT model and the actual content transfer is modeled by a probabilistic copying process similar to the SI epidemic model. While the HILT model evolves independently, the evolution of the SI model depends upon the HILT model, since the copying process takes into account the relay/destination state of the receiving node, indicated by the *want* bit of the node. In this section we will derive an o.d.e. limit for the SI part of the joint process. We shall also assume for simplicity that the thresholds in the HILT model are uniformly distributed; the analysis can be easily extended to HILT with general distribution F .

Pairwise meetings of the nodes constitute points of a Poisson process with rate λ , and let $\alpha_N = \frac{\alpha}{N}$ and $\sigma_N = \frac{\sigma}{N}$ be the copy probabilities to the destinations (nodes that are interested in the content) and relays (nodes not yet interested in the content) respectively. Recall from Section II-C (Figure 3) $\mathcal{X}(k)$ the set of destinations that have the content and $\mathcal{Y}(k)$ the set of relays that have the content with $X(k)$ and $Y(k)$ the sizes of these sets respectively. Taking into account the HILT driven conversion of relays that have the content into destinations, we note that both X_k and Y_k can be written in terms of the binomial random variables $P_{(\cdot)}(k)$ and $Q_{xy}(k)$, the first due to the content copying process (SI model) and the second due to the interest evolution process (HILT model). We can write,

$$X(k+1) = X(k) + P_x(k) + Q_{xy}(k)$$

$$Y(k+1) = Y(k) + P_y(k) - Q_{xy}(k)$$

Since the nodes meet at rate λ over a given slot, the probability that a destination node without the content, receives the content is $1 - e^{-\lambda\alpha_N(X(k)+Y(k))}$. Similarly for a relay node, the probability would be $1 - e^{-\lambda\sigma_N(X(k)+Y(k))}$. For large N , we can then write,

$$P_x(k) \stackrel{\text{dist.}}{=} \text{Bin}\left(A(k) - X(k), \lambda\alpha_N(X(k) + Y(k))\right)$$

$$P_y(k) \stackrel{\text{dist.}}{=} \text{Bin}\left(N - A(k) - Y(k), \lambda\sigma_N(X(k) + Y(k))\right)$$

Note also that due to the Interest Evolution process (HILT model) some relays get converted into destinations. Hence nodes from \mathcal{Y}_k (relays that have the content) may transition to \mathcal{X}_k (destinations that have the content). Since we are working with the uniform distribution, the probability of such a transition for each of the relay nodes would be $\frac{\gamma_N D(k)}{1 - \gamma_N B(k)}$. We denote by $Q_{xy}(k)$ the number of relay nodes that get converted during the k th time slot. Then,

$$Q_{xy}(k) \stackrel{\text{dist.}}{=} \text{Bin}\left(Y(k), \frac{\gamma_N D(k)}{1 - \gamma_N B(k)}\right)$$

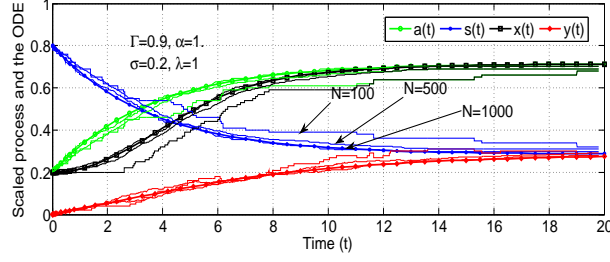


Fig. 8. Comparison of the scaled HILT-SI process for $N = 50, 100, 1000$ with the corresponding fluid limits.

Hence we can write the evolution of the processes $X(k)$ and $Y(k)$ as follows.

$$\begin{aligned} X(k+1) - X(k) &= \lambda \alpha_N (X(k) + Y(k)) (A(k) - X(k)) \\ &\quad + \frac{\gamma_N D(k)}{1 - \gamma_N B(k)} Y(k) + Z_x(k) \end{aligned}$$

$$\begin{aligned} Y(k+1) - Y(k) &= \lambda \sigma_N (X(k) + Y(k)) (N - A(k) - Y(k)) \\ &\quad - \frac{\gamma_N D(k)}{1 - \gamma_N B(k)} Y(k) + Z_y(k) \end{aligned}$$

where $Z_x(k)$ and $Z_y(k)$ are zero mean noise variables respectively for the X_k and Y_k processes respectively. In order to obtain the fluid limit, we work with the scaled process which evolves over mini-slots of width $\frac{1}{N}$. In each mini slot, the probability that a destination node without the content, receives the content will then be $1 - e^{-\frac{\lambda \alpha_N}{N} (X(k) + Y(k))}$ and a similar scaling occurs for the relay case. The interest evolution scaling in the $Q_{xy}(k)$ term is similar to the probabilistic scaling adopted in Section III-A. We can then write the scaled processes $X^N(t)$ and $Y^N(t)$ and the corresponding fractional processes $\tilde{X}^N(t)$ and $\tilde{Y}^N(t)$ (see Appendix D for details) and state the following theorem, along similar lines to Theorem 1.

Theorem 4: Given the joint evolution Markov process $(\tilde{B}^N(t), \tilde{D}^N(t), \tilde{X}^N(t), \tilde{Y}^N(t))$, we have for each $T > 0$ and each $\epsilon > 0$,

$$\mathbb{P} \left(\sup_{0 \leq u \leq T} \left| (\tilde{B}^N(\lfloor Nu \rfloor), \tilde{D}^N(\lfloor Nu \rfloor), \tilde{X}^N(\lfloor Nu \rfloor), \tilde{Y}^N(\lfloor Nu \rfloor)) - (b(u), d(u), x(u), y(u)) \right| > \epsilon \right) \xrightarrow{N \rightarrow \infty} 0$$

where $(b(u), d(u), x(u), y(u))$ is the unique solution to the o.d.e.,

$$\dot{b} = d \tag{9}$$

$$\dot{d} = \frac{\Gamma d}{1 - \Gamma b} (1 - b - d) - d \tag{10}$$

$$\dot{x} = \lambda \alpha (x + y) (a - x) + \frac{\Gamma d}{1 - \Gamma b} y \tag{11}$$

$$\dot{y} = \lambda \sigma (x + y) (1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y \tag{12}$$

with initial conditions $(b(0) = 0, d(0) = d_0, x(0) = x_0, y(0) = y_0)$.

Proof: A detailed proof verifying the conditions for Kurtz's theorem to hold, are presented in Appendix F-B. ■

A. Accuracy of the Fluid limit

Figure 8 shows the convergence of the scaled HILT-SI $(\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor), \tilde{X}^N(\lfloor Nt \rfloor), \tilde{Y}^N(\lfloor Nt \rfloor))$ process to the solutions of the above o.d.e $(b(t), d(t), x(t), y(t))$ for increasing values of $N = 100, 500, 1000$. We see that for $N = 1000$ the o.d.e approximates the scaled HILT-SI Markov chain very well.

The original HILT-SI process $(B(k), D(k), X(k), Y(k))$ is then compared with the solutions of the o.d.e by superimposing multiple sample paths of the original discrete time HILT-SI process on the o.d.e solutions. (see Figure 9). We find that the o.d.e solution approximates the original process really well, and permits using the fluid limit approximation for sufficiently large N .

V. NUMERICAL RESULTS

In this section we discuss the applications of the fluid limits through numerical results for various optimization problems of practical interest. While the first part of the section deals with optimizing for the interest evolution process considered in isolation, the latter part deals with optimizations for the joint evolution process. As an example, we work with HILT under uniform threshold distribution in this section.

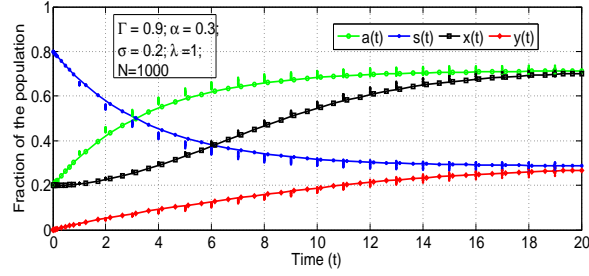


Fig. 9. Comparison of the unscaled HILT-SI process ($N = 1000$) with the corresponding fluid limit.

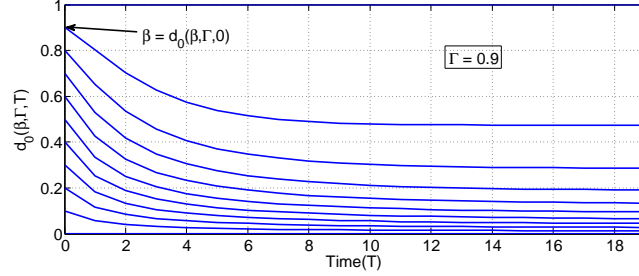


Fig. 10. Variation of d_0^* versus T for various values of β with $\Gamma = 0.9$

A. Interest Evolution

Content creators are often interested in understanding the evolution of popularity of their content, and would wish to maximize the level of popularity achieved. This, in our model, is equivalent to the final fraction of destinations (nodes that are interested in receiving the content). Once the content is created, in most cases, the content creator does not have control over influence weight Γ or the threshold distribution F of the population. In order to increase the spread of interest in the content, the only parameter that can be controlled is d_0 , the initial fraction of destinations in the population. In Section III, we derived the relation between d_0 and b_∞ , the final fraction of destinations. We might be interested in choosing the right d_0 which can give us the required b_∞ , and we see that by rearranging Equation (8) we get,

$$d_0 = \frac{b_\infty(1 - \Gamma)}{1 - b_\infty\Gamma} \quad (13)$$

Since the o.d.e. provides a good approximation for the temporal evolution of interest, we can also obtain results for time-constrained spread of popularity, i.e., the time taken by the process for the spread of influence is also a constraint, in addition to the initial fraction of destinations.

Theorem 5: Given the initial fraction of destinations d_0 in an HILT network with parameter Γ , the time we have to wait to get the final fraction of destinations to be at least β ($\beta < \frac{d_0}{r}$) is given by,

$$T(\beta, d_0, \Gamma) = \frac{1}{r} \ln \left(\frac{1 - r}{1 - \frac{\beta}{d_0} r} \right)$$

where $r = 1 - \Gamma + \Gamma d_0$.

Proof: Firstly, note that since $a_\infty = b_\infty = \frac{d_0}{r}$, we cannot reach $\beta > \frac{d_0}{r}$. Since we are observing the process at a finite time T , $d(T)$ is not zero. Hence, we consider $a(T) = b(T) + d(T)$ and set it to β . We get,

$$a(T) = d_0 \left(\frac{1}{r} - \left(\frac{1}{r} - 1 \right) e^{-rT} \right) = \beta \quad (14)$$

Rearranging terms, we get the expression for $T(\beta, d_0, \Gamma)$. ■

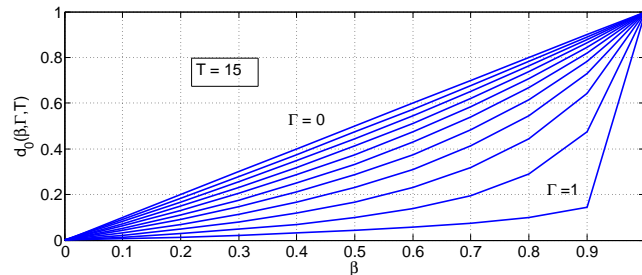


Fig. 11. Variation of d_0^* versus β for various values of Γ at $T = 15$

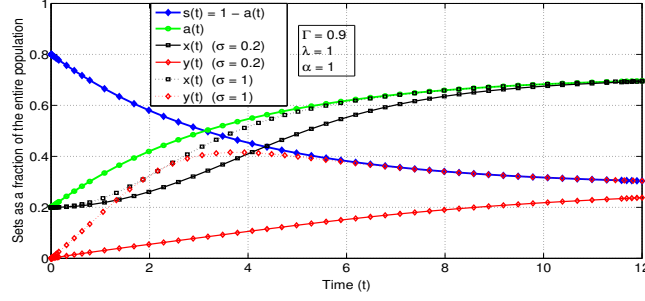


Fig. 12. Numerical simulation of HILT-SI model for two different values of relay copy probability $\sigma = 1$ and $\sigma = 0.2$.

Another interesting question would be to determine the initial fraction d_0 to be chosen so that by time T we will have at least β fraction of the nodes in the destination set, in the HILT network with parameter Γ . This can be solved numerically using the following fixed point equation obtained from Equation (14).

$$e^{-rT} = \frac{1 - \frac{\beta}{d_0}r}{1 - r}$$

Let $H(d_0) = e^{-rT}$ and $G(d_0) = \frac{1 - \frac{\beta}{d_0}r}{1 - r}$. We know that d_0^* that solves $H(d_0) = G(d_0)$ will lie in $[\frac{\beta(1-\Gamma)}{1-\beta\Gamma}, 1]$ and that the solution is unique, since $a(T)$ is a monotonic function in d_0 . We also know that for $d_0 < d_0^*$, $H(d_0) > G(d_0)$ and for $d_0 > d_0^*$, $H(d_0) < G(d_0)$. Under the above conditions, we can use the iterative bisection method that will converge to d_0^* .

The variation of d_0^* with respect to the parameters β , Γ and T can be seen in Figures 10 and 11. It is clear that,

- For fixed Γ and T , a higher value of β requires a higher d_0 (Figure 11).
- If the network has a high Γ , then it is sufficient to start with a smaller d_0 to reach a given β by T . (Figure 11).
- As the fixed time of interest T is increased (the time constraint is relaxed), the required initial d_0 decreases, from β at $T = 0$ and finally reaches d_0 obtained by setting $b_\infty = \beta$ in Equation (13) (Figure 10).

B. Joint evolution of interest and spread

In this section we discuss the optimization problems that might be of interest for the joint evolution process. Given that the network parameters Γ and d_0 are fixed, interest in the content evolves independently. The main motive of the content provider would then be to ensure that the content is delivered to as many destinations as possible.

Recall the o.d.e.s for the HILT-SI model given by Equations 9-12. Since we intend to deliver to the destinations, it would be optimal to set $\alpha = 1$. Hence the only control parameter is the copy probability to a relay, σ . This may be controlled by incentivizing or penalizing copying to relay nodes. As noted earlier, copying to a relay has two advantages, since the relay might meet a destination in the future, or might itself get converted into one, by the HILT process. But having a high σ value, will lead to increase in the number of *wasted* copies, the number of relays (not interested) that will end up having the content. The joint evolutions o.d.e.s have been simulated numerically for $\alpha = 1$ and two different values of σ and are plotted in Figure 12. As discussed, having a higher value of σ accelerates the $x(t)$ process, but leads to higher value of $y(t)$. Hence there is a tradeoff between achieving a high $x(t)$ while keeping $y(t)$ under bounds. We discuss two possible optimization problems for σ that can be posed, keeping in mind this tradeoff between $x(t)$ and $y(t)$. To understand clearly the effect of σ on the copying process, we work with $\Gamma = 0.9$ and have fixed the initial fraction of destinations $d_0 = 0.2$, initial fraction of destinations that have the content $x_0 = 0.2$. This implies that the set of nodes initially interested in the content is equal to the set of nodes that have the content. We assume that the content is delivered a priori to these initial destinations by some other means.

1) *Maximize target spread*: As content providers, we might be interested to deliver the content to as many destinations as possible by some fixed time. This might be the case, when the content is time-dependent and its usefulness expires by that fixed time. The fraction of destinations is itself increasing with time, due to the HILT interest evolution process. We would also wish to ensure that the number of relays that have the content, by that fixed time, is minimal. Our aim is then to maximize the value of $x(t)$ at some given fixed τ , subject to keeping $y(\tau)$ under control. The problem can be formally posed as,

$$\max_{\{\sigma: y(\tau) \leq \zeta\}} x(\tau)$$

Numerically, we can compute the feasible set $\{\sigma : y(\tau) \leq \zeta\}$ obtain the optimal solution σ^* that maximizes $x(\tau)$. We also observe experimentally that for a fixed τ , $x(\tau)$ and $y(\tau)$ are monotonically increasing with σ .

Figure 13 shows the variation of σ^* and the corresponding optimal $x(\tau)^*$ for various values of τ and a fixed fraction ζ . Observe that, for τ small, we can afford a high value of σ since $y(t)$ cannot exceed ζ in such a short time. Also, when τ is large, we can have a high σ , provided the interest evolution process is strong enough (high Γ). This will ensure that the number of relays $s(t)$ will be low enough by τ , and since $y(t) \leq s(t)$, $\forall t$, $y(t)$ cannot exceed ζ .

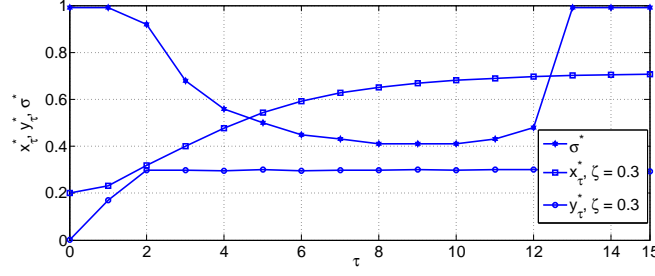


Fig. 13. Maximize target spread: The optimal solution plotted for a fixed value of ζ and varying τ .

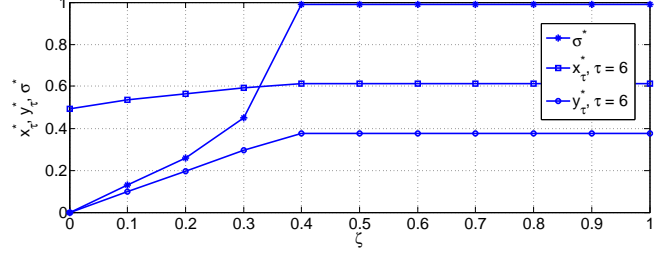


Fig. 14. Maximize target spread: The optimal solution plotted for a fixed value of τ and varying ζ .

Figure 14 shows the similar variations for various fractions ζ and a fixed time τ . As ζ , the constraint on the number of relay nodes with the content at τ is relaxed, we see that the optimal σ^* increases. It can be seen that $y(\tau)$ stabilizes at the value dictated by the number of relay nodes $s(\tau)$ for high value of σ . Also note that, if the fixed time of interest is high, then there is negligible contribution due to increase in σ since $x(t)$ increases to $a(t)$ eventually irrespective of σ (see Figure 12).

2) *Minimize Reach time*: Another problem of interest is to deliver the content to a fraction of the destinations as early as possible. This might be the case when, delivering to a considerable fraction of destinations accrues some benefit for the content provider from the content creator, and would like to earn the benefit as early as possible. Hence we try to minimize the time taken to reach a particular value of $x(t)$. Define $\tau_\eta = \inf\{t : x(t) \geq \eta\}$. Then we can formally pose this problem as,

$$\min_{\{\sigma : y(\tau_\eta) \leq \zeta\}} \tau_\eta$$

We numerically compute the feasible set by computing τ_η for each σ and verifying if $y(\tau_\eta) \leq \zeta$. We experimentally observe that for fixed η , τ_η is monotonic decreasing in σ and $y(\tau_\eta)$ is monotonic increasing in σ .

Figures 15 and 16 show the numerically evaluated optimal solutions for the above optimization problem, for various values of η and ζ . In Figure 15, it is clear that a higher value of ζ allows us to use a high value of σ so that we can accelerate the content delivery to destinations. As a result, we also see that the time taken to deliver the content decreases with increasing ζ .

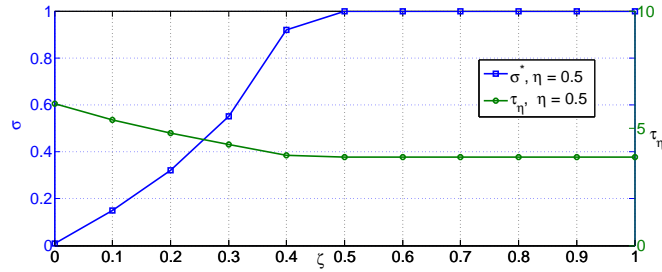


Fig. 15. Minimize reach time: The optimal solution plotted for a fixed value of η and varying ζ .

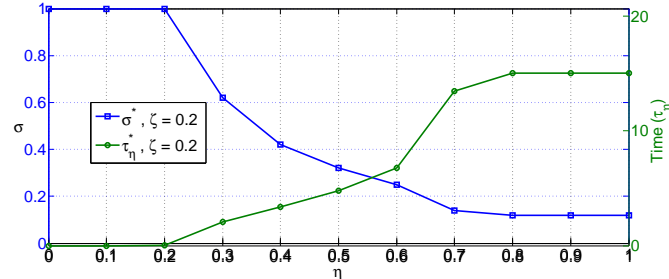


Fig. 16. Minimize reach time: The optimal solution plotted for a fixed value of ζ and varying η .

In Figure 16, note that since we begin with $x_0 = 0.2$, for $\eta \leq 0.2$, $\tau_\eta = 0$ and σ can be any value. When η increases, as long as τ_η is low enough, we can afford to have a high σ and still keep $y(\tau_\eta)$ under control, since the $y(t)$ process may not have enough time to reach ζ . But as η increases further, σ needs to be reduced, since an increased value of τ_η implies, the process will run longer and $y(\tau_\eta)$ might then exceed ζ .

VI. CONCLUSIONS

In this paper we studied the joint evolution of popularity and spread of content in a mobile P2P environment. We used two different models to capture the processes of interest evolution (HILT model) and content delivery (SI model), and derive their fluid limits. We then showed that the classic SIR epidemic model can be derived as a special case of the HILT model and derived explicit solutions for the model when the thresholds were uniformly distributed. We then used the fluid limits to address several optimization problems that might be of practical interest. This work can be extended in several possible directions. The co-evolution problem can be studied for other related optimization problems, and we can try obtaining explicit optimal solutions for a few special cases. One could generalize the problem for analysing the spread of popularity and spread for multiple P2P content. And finally one could consider other models for both interest evolution and for the copying process, including those where the processes of interest evolution and content delivery are dependent on each other.

REFERENCES

- [1] A. Diaz, P. Merino, L. Panizo, and A. M. Recio, "A survey on mobile peer-to-peer technology," in *JCS*, Nov. 2007, pp. 59–68.
- [2] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *SIGKDD*, 2002.
- [3] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *ACM SIGKDD*, 2003.
- [4] M. H. R. Khoushani, S. Sarkar, and E. Altman, "Dispatch then stop: Optimal dissemination of security patches in mobile wireless networks," in *CDC*, 2010.
- [5] C. Singh, A. Kumar, R. Sundaresan, and E. Altman, "Optimal forwarding in delay tolerant networks with multiple destinations," in *WiOpt*, 2011.
- [6] T. Repantis and V. Kalogeraki, "Data dissemination in mobile p2p networks," in *MDM*, 2005.
- [7] S. Shakkottai and R. Johari, "Demand-aware content distribution on the internet," in *Transactions on Networking*, 2010.
- [8] W. O. Kermack and A. G. McKendrick, "A contribution the mathematical theory of epidemics," in *Proceedings of the Royal Society of London*, 1927.
- [9] R. Nunez-Queija and B. Prabhu, "Scaling laws for file dissemination in p2p networks with random contacts," in *IWQOS*, 2008.
- [10] L. J. Allen, "Some discrete time si,sir and sis epidemic models," in *Mathematical BioSciences*, 1994.
- [11] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump markov processes," in *Journal of Applied Probability*, 1970.
- [12] R. Darling, "Fluid limits of pure jump markov processes: A practical guide," 2002, available at arxiv.org/pdf/math/0210109.
- [13] R. Darling and J. Norris, "Differential equation approximations for markov chains," in *Probability Surveys*, vol. 5, 2008.
- [14] M. Benaïm and J.-Y. L. Boudec, "A class of mean field interaction models for computer and communication systems," in *Performance Evaluation*, 2008.
- [15] D. R. Cox, *Renewal Theory*, M. S. Bartlett, Ed. Metheun & Co. Ltd. Science Paperbacks, 1961.
- [16] M. Granovetter, "Threshold models of collective behaviour," in *The American Journal of Sociology*, 1978.
- [17] S. Venkatramanan and A. Kumar, "Information dissemination in social networks under the linear threshold model," in *National Conference on Communications*, 2011.
- [18] D. J. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge University Press, 1999.

APPENDIX A

INTEREST EVOLUTION IS A MARKOV CHAIN

Proposition 1: For the HILT model, $(B(k), D(k)), k \geq 0$, is a discrete time Markov chain (DTMC).

Proof: Since, for each $k \geq 0$, $B(k+1) = B(k) + D(k)$, it suffices to show that, for $k \geq 0$,

$$P(D(k+1) = \ell | (B(0), D(0)), (B(1), D(1)), \dots, ((B(k), D(k)) = (j, m)))$$

is a function only of (j, m) . We need the following simple lemma (in which the new notation is local to the lemma).

Lemma 1: Given (X_1, X_2, \dots, X_n) i.i.d. random variables with common c.d.f. $G(x), x \geq a$, where $a > 0$, and given $b > a$, define

$$\mathcal{Z} = \{i : 1 \leq i \leq n, X_i > b\}, \text{ and } Z = |\mathcal{Z}|.$$

Let Y_1, Y_2, \dots, Y_Z be random variables jointly distributed as $X_i, i \in \mathcal{Z}$ (ordered in the increasing sequence of indices). Then

$$P(Y_i \leq y_i, 1 \leq i \leq Z | Z = m) = \prod_{i=1}^m \frac{G(y_i) - G(b)}{1 - G(b)}$$

for $y_i \geq b, 1 \leq i \leq m$.

Proof: (of Lemma 1) It is easily seen that

$$P(Z = m, Y_1 \leq y_1, \dots, Y_m \leq y_m) = \binom{n}{m} (G(b))^{n-m} \prod_{i=1}^m (G(y_i) - G(b))$$

and

$$P(Z = m) = \binom{n}{m} (G(b))^{n-m} (1 - G(b))^m$$

from which the desired result immediately follows. ■

Returning to the proof of Proposition 1, given $(B(0), D(0)), (B(1), D(1)), \dots, ((B(k), D(k)) = (j, m))$, via a recursive application of Lemma 1 (starting from the initial i.i.d. thresholds $\Theta_i, 1 \leq i \leq N$, with common c.d.f. $F(\theta)$), we conclude that the thresholds of the users in $\mathcal{N} \setminus (\mathcal{B}(k) \cup \mathcal{D}(k))$ are i.i.d. with common c.d.f. $\frac{F(\theta) - F(\gamma j)}{1 - F(\gamma j)}$, over the range $\theta \geq \gamma j$. At the end of period k , the newly interested users in $\mathcal{D}(k)$ will serve as additional influence on the users in $\mathcal{N} \setminus (\mathcal{B}(k) \cup \mathcal{D}(k))$. Of these, ℓ will become interested (i.e., $D(k+1) = \ell$) with probability

$$\binom{(N - (j + m))}{\ell} \left(\frac{F(\gamma(j + m)) - F(\gamma j)}{1 - F(\gamma j)} \right)^\ell \left(1 - \frac{F(\gamma(j + m)) - F(\gamma j)}{1 - F(\gamma j)} \right)^{(N - (j + m)) - \ell}$$

which depends on the “history” only via (j, m) , and thereby establishes the desired result. ■

APPENDIX B FLUID LIMIT FOR THE HILT MODEL

Recalling from Section III-A, the evolution of $(\tilde{B}^N(t), \tilde{D}^N(t))$ can be written in terms of its mean “drift” at any t as follows.

$$\begin{aligned} \tilde{B}^N(t+1) &= \tilde{B}^N(t) + \frac{\tilde{D}^N(t)}{N} + \tilde{Z}_b^N(t+1) \\ \tilde{D}^N(t+1) &= \frac{N-1}{N} \tilde{D}^N(t) \\ &+ \mathbb{E} \left[\frac{F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma_N N(\tilde{B}^N(t)))}{1 - F(\gamma_N N(\tilde{B}^N(t)))} \right] \times \\ &\quad (1 - \tilde{B}^N(t) - \tilde{D}^N(t)) + \tilde{Z}_d^N(t+1) \end{aligned}$$

where we have defined $\tilde{Z}_b^N(t)$ and $\tilde{Z}_d^N(t)$ in an analogous manner. The mean drift rate function, $g^N(\tilde{B}^N(k), \tilde{D}^N(k))$ per unit update step size $\frac{1}{N}$ becomes,

$$\frac{g^N(\tilde{B}^N(t), \tilde{D}^N(t))}{\frac{1}{N}} = \left(g_1^N(\cdot), g_2^N(\cdot) \right)$$

where

$$\begin{aligned} g_1^N(\cdot) &= \tilde{D}^N(t) \\ g_2^N(\cdot) &= N \mathbb{E} \left[\frac{F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma_N N(\tilde{B}^N(t)))}{1 - F(\gamma_N N(\tilde{B}^N(t)))} \right] \times \\ &\quad (1 - \tilde{B}^N(t) - \tilde{D}^N(t)) - \tilde{D}^N(t) \end{aligned}$$

Recall $\gamma_N \times (N-1) = \Gamma$ and $f(x)$ the density function of the threshold distribution F . We can then show that (see Appendix C),

$$\begin{aligned} \lim_{N \rightarrow \infty} N \mathbb{E} \left[\frac{F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma_N N(\tilde{B}^N(t)))}{1 - F(\gamma_N N(\tilde{B}^N(t)))} \right] \\ = \frac{f(\Gamma b) \Gamma d}{1 - F(\Gamma b)} \end{aligned}$$

Letting $g_1(b, d) = d$ and $g_2(b, d) = \frac{f(\Gamma b) \Gamma d}{1 - F(\Gamma b)} (1 - b - d) - d$ and define,

$$g(b, d) := \left(g_1(b, d), g_2(b, d) \right)$$

We can then use these limiting drift functions to obtain the fluid limits corresponding to the HILT model.

APPENDIX C

Consider

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[\frac{F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma_N N(\tilde{B}^N(t)))}{1 - F(\gamma_N N(\tilde{B}^N(t)))} \right]$$

where the expectation is with respect to $\tilde{C}^N(t)$ given $(\tilde{B}^N(t), \tilde{D}^N(t))$. Applying Taylor’s expansion to $F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t)))$ around $\tilde{B}^N(t)$ we get,

$$\begin{aligned}
& \lim_{N \rightarrow \infty} N \mathbb{E} \left[\frac{F(\gamma N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma N(\tilde{B}^N(t)))}{1 - F(\gamma N(\tilde{B}^N(t)))} \right] \\
&= \lim_{N \rightarrow \infty} N \mathbb{E} \left[\frac{\gamma N \tilde{C}^N(t) f(\gamma N \tilde{B}^N(t))}{1 - F(\gamma N(\tilde{B}^N(t)))} \right] \\
&+ \lim_{N \rightarrow \infty} N \mathbb{E} \left[\frac{\gamma N \frac{\tilde{C}^N(t)^2}{2} \dot{f}(\zeta)}{1 - F(\gamma N(\tilde{B}^N(t)))} \right]
\end{aligned}$$

In the second term since $\tilde{C}^N(t) = \frac{C^N(t)}{N}$ where $C^N(t) \sim \text{Bin}(D^N(t), \frac{1}{N})$, we have

$$\begin{aligned}
\mathbb{E}(\tilde{C}^N(t)^2) &= \frac{1}{N^2} \mathbb{E}(C^N(t)^2) \\
&= \frac{1}{N^2} \left[D^N(t) \frac{1}{N} (1 - \frac{1}{N}) + (\frac{D^N(t)}{N})^2 \right] \\
&= \frac{1}{N^2} \left[\tilde{D}^N(t) (1 - \frac{1}{N}) + (\tilde{D}^N(t))^2 \right]
\end{aligned}$$

Also $\dot{f}(\zeta)$ is bounded, since we require the drift functions to satisfy the Lipschitz condition. Hence the second term vanishes as $N \rightarrow \infty$. In the first term, noting that $\gamma N \rightarrow \Gamma$ and $\mathbb{E}(\tilde{C}^N(t)) = \frac{\tilde{D}^N(t)}{N}$, given $\tilde{B}^N(t) = b$, $\tilde{D}^N(t) = d$, the limit becomes $\frac{\Gamma d f(\Gamma b)}{1 - F(\Gamma b)}$.

APPENDIX D FLUID LIMIT FOR THE HILT-SI MODEL

Adopting the scaling as explained in Section IV, we can write the evolution of $(X^N(t), Y^N(t))$ as follows:

$$\begin{aligned}
& X^N(t+1) - X^N(t) \\
&= \lambda \alpha_N \frac{(X^N(t) + Y^N(t))}{N} (A^N(t) - X^N(t)) \\
&+ \frac{\gamma_N \frac{D^N(t)}{N}}{1 - \gamma_N \tilde{B}^N(t)} Y^N(t) + Z_x^N(t) \\
& Y^N(t+1) - Y^N(t) \\
&= \lambda \sigma_N \frac{(X^N(t) + Y^N(t))}{N} (N - A^N(t) - Y^N(t)) \\
&- \frac{\gamma_N \frac{D^N(t)}{N}}{1 - \gamma_N \tilde{B}^N(t)} Y^N(t) + Z_y^N(t)
\end{aligned}$$

As earlier, defining $\tilde{X}^N(t)$, $\tilde{Y}^N(t)$, $\tilde{A}^N(t)$, $\tilde{B}^N(t)$, $\tilde{D}^N(t)$ as the fractions of entire population, we can write the mean drift functions per unit update size for $\tilde{X}^N(t)$ and $\tilde{Y}^N(t)$ respectively as,

$$\begin{aligned}
\frac{g_x^N(x, y, b, d)}{\frac{1}{N}} &= \left(\lambda \alpha_N N (x + y) (a - x) + \frac{\gamma_N N d}{1 - \gamma_N N b} y \right) \\
\frac{g_y^N(x, y, b, d)}{\frac{1}{N}} &= \left(\lambda \sigma_N N (x + y) (1 - a - y) - \frac{\gamma_N N d}{1 - \gamma_N N b} y \right)
\end{aligned}$$

Let $g_x(x, y, b, d) = \lambda \alpha (x + y) (a - x) + \frac{\Gamma d}{1 - \Gamma b} y$ and $g_y(x, y, b, d) = \lambda \sigma (x + y) (1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y$. Combining these with the limiting drift functions of the HILT model (see Appendix B) we can derive the fluid limits of the HILT-SI model.

APPENDIX E KURTZ'S THEOREM AND PROOF OF THE THEOREM 1

Kurtz's theorem [11] provides us a way by which we can, subject to certain conditions, approximate the evolution of a pure jump Markov process by the solution of a derived ODE. In this section, we derive the necessary conditions for the convergence

of the pure jump Markov process to the ODE solution, along the lines of Kurtz theorem, and subsequently check them for the HILT model.

For each $N \geq 1$, $Y^{(N)}(k), k \geq 0$, is a DTMC on $\Delta^{(N)} \subset \Delta$ where Δ is a subset of an appropriate dimensional Euclidean space. Let $e_j^{(N)}, 1 \leq j \leq J^{(N)}$, be the possible jumps of $Y^{(N)}(\cdot)$ and $p_j^{(N)}(y), y \in \Delta^{(N)}$, the jump probabilities, i.e.,

$$p_j^{(N)}(y) := P(Y^{(N)}(k+1) - Y^{(N)}(k) = e_j^{(N)} | Y^{(N)}(k) = y)$$

We define an indicator variable, $I_j(Y^{(N)}(k)) = 1$ if the jump out of $Y^{(N)}(k)$ is $e_j^{(N)}$. Then we can write

$$\begin{aligned} Y^{(N)}(k) &= Y^{(N)}(0) + \sum_{i=0}^{k-1} \sum_{j=1}^{J^{(N)}} e_j^{(N)} I_j(Y^{(N)}(i)) \\ &= Y^{(N)}(0) + \sum_{i=0}^{k-1} \sum_{j=1}^{J^{(N)}} e_j^{(N)} p_j^{(N)}(Y^{(N)}(i)) \\ &\quad + \sum_{i=0}^{k-1} \sum_{j=1}^{J^{(N)}} e_j^{(N)} (I_j(Y^{(N)}(i)) - p_j^{(N)}(Y^{(N)}(i))) \end{aligned}$$

where in the second equality, the second terms comprises the sum of the successive mean jumps, and the last term the sum of successive random "noise" terms. We rewrite these compactly by defining, for $i \geq 0$,

Define, for $i \geq 0$

$$\xi^{(N)}(i) := \sum_{j=1}^{J^{(N)}} e_j^{(N)} (I_j(Y^{(N)}(i)) - p_j^{(N)}(Y^{(N)}(i))),$$

so that

$$Y^{(N)}(k) = Y^{(N)}(0) + \sum_{i=0}^{k-1} f^{(N)}(Y^{(N)}(i)) + \sum_{i=0}^{k-1} \xi^{(N)}(i)$$

where $f^{(N)}(y) := \sum_{j=1}^{J^{(N)}} e_j^{(N)} p_j^{(N)}(y)$ is the expected drift out of the state $y \in \Delta^{(N)}$.

With $\mathcal{F}_k^{(N)} := \sigma(Y^{(N)}(0), Y^{(N)}(1), \dots, Y^{(N)}(k))$, we easily see that

Lemma 2: $Z^{(N)}(k) := \sum_{i=0}^{k-1} \xi^{(N)}(i)$, $k > 0$, is a martingale with respect to $\mathcal{F}_k^{(N)}$, $k \geq 0$.

For each $t \geq 0$, for each N , we can write

$$Y^{(N)}(\lfloor Nt \rfloor) \tag{15}$$

$$= Y^{(N)}(0) + \sum_{i=0}^{\lfloor Nt \rfloor - 1} f^{(N)}(Y^{(N)}(i)) + Z^{(N)}(\lfloor Nt \rfloor) \tag{16}$$

Suppose

$$\sup_{x \in \Delta^{(N)}} \left\| \frac{f^{(N)}(x)}{\frac{1}{N}} - f(x) \right\| \xrightarrow{N \rightarrow \infty} 0 \tag{17}$$

with $f(x)$ being Lipschitz on Δ , i.e., for all $x, y \in \Delta$, for some $L > 0$,

$$\|f(x) - f(y)\| \leq L\|x - y\| \tag{18}$$

Now, in (15), if $Y^{(N)}(0) \xrightarrow{P} y(0)$, and the second term goes to 0, the process $Y^{(N)}(\lfloor Nt \rfloor), t \geq 0$, appears to be following

$$y(t) = y(0) + \int_0^t f(y(u)) du$$

Take $t \in [0, T]$. For given t , subtracting the ODE solution on both sides of (15)

$$Y^{(N)}(\lfloor Nt \rfloor) - y(t) \tag{19}$$

$$= Y^{(N)}(0) - y(0) + \sum_{i=0}^{\lfloor Nt \rfloor - 1} f^{(N)}(Y^{(N)}(i)) \tag{20}$$

$$- \int_0^t f(y(u)) du + Z^{(N)}(\lfloor Nt \rfloor) \tag{21}$$

We will study each of the three "error" terms on the right hand side of this equation. First, we consider the term

$$\sum_{i=0}^{\lfloor Nt \rfloor - 1} f^{(N)}(Y^{(N)}(i)) - \int_0^t f(y(u)) du \quad (22)$$

$$= \sum_{i=0}^{\lfloor Nt \rfloor - 1} f^{(N)}(Y^{(N)}(i)) - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f(Y^{(N)}(i)) \quad (23)$$

$$+ \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f(Y^{(N)}(i)) - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f\left(y\left(\frac{i}{N}\right)\right) \quad (24)$$

$$+ \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f\left(y\left(\frac{i}{N}\right)\right) - \int_0^t f(y(u)) du \quad (25)$$

Term (23): Using the uniform convergence hypothesis, for each t , and for an $\epsilon_1 > 0$ there exists an N large so that,

$$\begin{aligned} & \left\| \sum_{i=0}^{\lfloor Nt \rfloor - 1} f^{(N)}(Y^{(N)}(i)) - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f(Y^{(N)}(i)) \right\| \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \left\| \frac{f^{(N)}(Y^{(N)}(i))}{\frac{1}{N}} - f(Y^{(N)}(i)) \right\| \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \epsilon_1 = \frac{\lfloor Nt \rfloor}{N} \epsilon_1 \end{aligned}$$

Term (25): By the hypothesis in (17), $f(\cdot)$ is Lipschitz on Δ . The following can then be proved:

- 1) $\|f(y)\| \leq \|f(0)\| + L\|y\|$ (i.e., at most linear growth in $\|y\|$)
- 2) The ODE $\dot{y}(t) = f(y(t))$, with initial value $y(0)$, has the unique solution $y(t) = y(0) + \int_0^t f(y(u)) du$. Further, over $0 \leq t \leq T$, there are constants such that
 - a) $\|f(y(t))\| \leq K_1(T)(1 + \|y(0)\|)$
 - b) $\|y(t + \tau) - y(t)\| \leq K_1(T)(1 + \|y(0)\|)\tau$

$$\begin{aligned} & \left\| \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f\left(y\left(\frac{i}{N}\right)\right) - \int_0^t f(y(u)) du \right\| \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} \left\| \frac{1}{N} f\left(y\left(\frac{i}{N}\right)\right) - \int_{\frac{i}{N}}^{\frac{i+1}{N}} f(y(u)) du \right\| \\ & + \left\| \int_{\frac{\lfloor Nt \rfloor}{N}}^t f(y(u)) du \right\| \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{2N} \left\| \left[f\left(y\left(\frac{i+1}{N}\right)\right) - f\left(y\left(\frac{i}{N}\right)\right) \right] \right\| + c'_1(T) \frac{1}{N} \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} c_1(T) \frac{1}{2N^2} + c'_1(T) \frac{1}{N} \\ & = c_1(T) \frac{\lfloor Nt \rfloor}{2N^2} + c'_1(T) \frac{1}{N} \end{aligned}$$

where the penultimate inequality follows from the Lipschitz property of $f(\cdot)$.

Term (24): By using the Lipschitz property of $f(\cdot)$

$$\begin{aligned} & \left\| \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f(Y^{(N)}(i)) - \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} f\left(y\left(\frac{i}{N}\right)\right) \right\| \\ & \leq \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} L \left\| Y^{(N)}(i) - y\left(\frac{i}{N}\right) \right\| \end{aligned}$$

We now work on the left hand side of (19)

$$\begin{aligned}
& \left\| Y^{(N)}(\lfloor Nt \rfloor) - y(t) \right\| \\
&= \left\| Y^{(N)}(\lfloor Nt \rfloor) - y\left(\frac{\lfloor Nt \rfloor}{N}\right) - \left(y(t) - y\left(\frac{\lfloor Nt \rfloor}{N}\right)\right) \right\| \\
&\geq \left\| Y^{(N)}(\lfloor Nt \rfloor) - y\left(\frac{\lfloor Nt \rfloor}{N}\right) \right\| - \left\| y(t) - y\left(\frac{\lfloor Nt \rfloor}{N}\right) \right\|
\end{aligned} \tag{26}$$

and it can be shown that, with $y(t) = y(0) + \int_0^t f(y(u))du$, for $t \in [0, T]$,

$$\left\| y(t) - y\left(\frac{\lfloor Nt \rfloor}{N}\right) \right\| \leq c_4(T) \frac{1}{N} \tag{27}$$

We now use the hypothesis that $Y^{(N)}(0) \xrightarrow{P} y(0)$

Then, for large enough N , combining (23), (25), (24), (26), (27) with (19) for all $t \in [0, T]$, we can write

$$\begin{aligned}
& \left\| Y^{(N)}(\lfloor Nt \rfloor) - y\left(\frac{\lfloor Nt \rfloor}{N}\right) \right\| \\
&\leq \epsilon + \sum_{i=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} L \left\| Y^{(N)}(i) - y\left(\frac{i}{N}\right) \right\| \\
&\quad + \left\| Z^{(N)}(\lfloor Nt \rfloor) \right\|
\end{aligned} \tag{28}$$

We now work on the "noise" martingale, using Burkholder Inequality(BI) for zero-mean martingales,

$$\begin{aligned}
& \mathbb{E} \left(\sup_{0 \leq t \leq T} \left\| Z^{(N)}(\lfloor Nt \rfloor) \right\|^2 \right) \\
&= \mathbb{E} \left(\sup_{0 \leq t \leq T} \left\| \sum_{i=0}^{\lfloor Nt \rfloor - 1} \xi^{(N)}(i) \right\|^2 \right) \\
&\stackrel{BI}{\leq} c_1 \mathbb{E} \left(\sum_{i=0}^{\lfloor NT \rfloor - 1} \left\| \xi^{(N)}(i) \right\|^2 \right) \\
&\leq c_1 c_0 \frac{\lfloor NT \rfloor}{N^2}
\end{aligned}$$

Where we have *required* that $\mathbb{E} \left\| \xi^{(N)}(i) \right\|^2 \leq c_0 \frac{1}{N^2}$, for all i .

It follows that

$$P \left(\sup_{0 \leq t \leq T} \left\| Z^{(N)}(\lfloor Nt \rfloor) \right\| \leq \epsilon \right) \xrightarrow{N \rightarrow \infty} 1$$

Theorem 6: (Gronwall's Inequality) Given a positive sequence $a_n, n \geq 0$, and $b, c \geq 0$, if, $x_n, n \geq 0$, is a nonnegative sequence such that, for each $k \geq 1$

$$x_k \leq b + c \sum_{i=0}^{k-1} a_i x_i$$

then, for each $k \geq 1$

$$x_k \leq b e^{c \sum_{i=0}^{k-1} a_i}$$

From 28, on a sample path where $\sup_{0 \leq t \leq T} \left\| Z^{(N)}(\lfloor Nt \rfloor) \right\| \leq \epsilon$ and $\|Y^{(N)}(0) - y(0)\| < \epsilon$, we can write, for $0 \leq k \leq \lfloor NT \rfloor$,

$$\left\| Y^{(N)}(k) - y\left(\frac{k}{N}\right) \right\| \leq 3\epsilon + \sum_{i=0}^{k-1} \frac{1}{N} L \left\| Y^{(N)}(i) - y\left(\frac{i}{N}\right) \right\|$$

Hence, by Gronwall's inequality, for any $\epsilon > 0$,

$$\begin{aligned}
& \mathbb{P} \left(\sup_{0 \leq t \leq T} \left\| Z^{(N)}(\lfloor Nt \rfloor) \right\| \leq \epsilon \right) \\
&\leq \mathbb{P} \left(\sup_{0 \leq k \leq \lfloor NT \rfloor} \left\| Y^{(N)}(k) - y\left(\frac{k}{N}\right) \right\| \leq 2\epsilon e^{\frac{L \lfloor NT \rfloor}{N}} \right)
\end{aligned}$$

By the earlier assertion, the left hand term $\rightarrow 1$ as $N \rightarrow \infty$, hence so does the right hand term. Hence we can state the following theorem.

Theorem 7: With the notation defined earlier, if, for all $k \geq 0$,

$$\mathbb{E} \left(\left\| \sum_{j=1}^{J^{(N)}} e_j^{(N)}(I_j(Y^{(N)}(k))) - f^{(N)}(Y^{(N)}(k)) \right\|^2 | \mathcal{F}_k \right) \leq c_0 \frac{1}{N^2}$$

Further, $\sup_{x \in \Delta^{(N)}} \left\| \frac{f^{(N)}(x)}{\frac{1}{N}} - f(x) \right\| \xrightarrow{N \rightarrow \infty} 0$ with $f(x)$ being Lipschitz (i.e., for all $x, y \in \Delta$, for some $L > 0$, $\|f(x) - f(y)\| \leq L \|x - y\|$). Define $y(t), 0 \leq t \leq T$, by $y(t) = y(0) + \int_0^t f(y(u))du$. Also, $Y^{(N)}(0) \xrightarrow{P} y(0)$.

Then, for each $T > 0$, and each $\epsilon > 0$,

$$P \left(\sup_{0 \leq t \leq T} \|Y^{(N)}(\lfloor Nt \rfloor) - y(t)\| > \epsilon \right) \xrightarrow{N \rightarrow \infty} 0$$

■

APPENDIX F

APPLICATION OF KURTZ'S THEOREM

A. Verification of Kurtz's theorem for HILT model

Now we can check each of the above conditions in Theorem 7 for the HILT model.

(i) Lipschitz property

Consider,

$$g_1(b, d) = d$$

$$g_2(b, d) = h_F(\Gamma b)\Gamma d(1 - b - d) - d$$

$$\frac{\partial g_1}{\partial b} = 0; \frac{\partial f_1}{\partial d} = 1$$

$$\frac{\partial g_2}{\partial b} = \Gamma^2 d \dot{h}_F(\Gamma b)(1 - b - d) - \Gamma d h_F(\Gamma b); \frac{\partial g_2}{\partial d} = h_f(\Gamma b)\Gamma(1 - b - 2d) - 1$$

Under the assumption that $\dot{f}(\cdot)$ is bounded, we see that each of the terms above is bounded when $(b, d) \in [0, 1] \times [0, 1 - b]$. Thus the norm of Jacobian $\|Dg(b, d)\|$ is uniformly bounded, and it follows that $g(b, d) = (g_1(b, d), g_2(b, d))$ is Lipschitz.

(ii) Uniform Convergence

We know that

$$g^N(\tilde{B}^N(k), \tilde{D}^N(k)) := \frac{1}{N} \left(\tilde{D}^N(k), N \mathbb{E} \left[\frac{F(\gamma_N N(\tilde{B}^N(t) + \tilde{C}^N(t))) - F(\gamma_N N(\tilde{B}^N(t)))}{1 - F(\gamma_N N(\tilde{B}^N(t)))} \right] \times \left(1 - \tilde{B}^N(t) - \tilde{D}^N(t) \right) - \tilde{D}^N(t) \right)$$

Let $g_1(b, d) = d$ and $g_2(b, d) = \frac{f(\Gamma b)\Gamma d}{1 - F(\Gamma b)}(1 - b - d) - d$ and define,

$$g(b, d) := \left(g_1(b, d), g_2(b, d) \right)$$

By definition, $\lim_{N \rightarrow \infty} N\gamma = \Gamma$ and by the steps in Appendix C, the uniform convergence of $\frac{g^N(b, d)}{1/N}$ to $g(b, d)$ in the domain $(b, d) \in [0, \frac{1}{N}, \dots, N] \times [0, \frac{1}{N}, \dots, b]$ is proven.

(iii) Bounded Noise variance

We will write the noise variance terms for the scaled processes $B^N(k)$ and $D^N(k)$, and from them derive the noise variance terms for the density dependent processes by scaling them by $\frac{1}{N^2}$.

Let $Z_b^N(k)$ and $Z_d^N(k)$ be the zero mean random variables representing the noise in the drift terms of $B^N(k)$ and $D^N(k)$ respectively. We can then write,

$$E[|Z_b^N(k)|^2 | \mathcal{F}_k] = (1 - \frac{1}{N}) \frac{1}{N} D^N(k)$$

$$E[|Z_d^N(k)|^2 | \mathcal{F}_k] = (1 - \frac{1}{N}) \frac{1}{N} D^N(k) + \sum_{z=0}^{z=D^N(k)} g(z)(1 - g(z))(N - B^N(k) - D^N(k)) \binom{D^N(k)}{z} \frac{1}{N}^z (1 - \frac{1}{N})^{D^N(k)-z}$$

where,

$$g(z) = \frac{F(\gamma N(B^N(k) + z)) - F(\gamma N B^N(k))}{1 - F(\gamma N B^N(k))}$$

Note that both these terms can be upper bounded by constants say c_b and c_d and hence the noise conditions for $\tilde{B}^N(k)$ and $\tilde{D}^N(k)$ are satisfied.

(iv) **Convergence of initial conditions** By choice, we have $\tilde{B}^N(0) = b(0)$ and $\tilde{D}^N(0) = d(0)$.

Thus by Theorem 7, we have for each $T > 0$ and each $\epsilon > 0$,

$$P\left(\sup_{0 \leq t \leq T} \|(\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor)) - (b(t), d(t))\| > \epsilon\right) \xrightarrow{N \rightarrow \infty} 0$$

where $(b(t), d(t))$ is the unique solution of the ODE

$$\begin{aligned} \dot{b}(t) &= d(t) \\ \dot{d}(t) &= -d(t) + \frac{\Gamma d(t)}{1 - \Gamma b(t)}(1 - b(t) - d(t)) \end{aligned}$$

with initial conditions $(b(0) = 0, d(0) = a(0))$. ■

B. Verification of Kurtz theorem for HILT-SI model

The first two equations, as proved above will remain unaltered. Consider the drift equations for $\tilde{X}^N(k)$ and $\tilde{Y}^N(k)$ and the corresponding ODEs. We can use a similar procedure as above to verify the conditions for applying Kurtz's theorem.

(i) Lipschitz property

Consider,

$$\begin{aligned} g_x(x, y, b, d) &= \lambda \alpha(x + y)(a - x) + \frac{\Gamma d}{1 - \Gamma b} y \\ g_y(x, y, b, d) &= \lambda \sigma(x + y)(1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y \end{aligned}$$

where $a = b + d$. We can calculate the partial derivatives of g_x and g_y with respect to x, y, b, d and note that given $b + d \leq 1, x \leq b + d, y \leq 1 - b - d$ and all quantities non-negative, we have that the Jacobian is uniformly bounded and hence the drift functions are Lipschitz.

(ii) Uniform Convergence

$$\begin{aligned} g_x^N(x, y, b, d) &= \frac{1}{N} \left(\lambda \alpha_N N(x + y)(a - x) + \frac{\gamma_N N d}{1 - \gamma_N b} y \right) \\ g_y^N(x, y, b, d) &= \frac{1}{N} \left(\lambda \sigma_N N(x + y)(1 - a - y) - \frac{\gamma_N N d}{1 - \gamma_N b} y \right) \end{aligned}$$

Define,

$$\begin{aligned} g_x(x, y, b, d) &= \lambda \alpha(x + y)(a - x) + \frac{\Gamma d}{1 - \Gamma b} y \\ g_y(x, y, b, d) &= \lambda \sigma(x + y)(1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y \end{aligned}$$

Since as $N \rightarrow \infty$, we have $\gamma_N N \rightarrow \Gamma, \alpha_N N \rightarrow \alpha, \sigma_N N \rightarrow \sigma$ and we see that the uniform convergence is straightforward.

(iii) Bounded Noise variance

As earlier we will write the noise variance terms for $X^N(k)$ and $Y^N(k)$ and derive the density dependent process' noise variance by scaling by $\frac{1}{N^2}$. Let $Z_x^N(k)$ and $Z_y^N(k)$ be the zero mean random variables representing the noise in the drift terms of $X^N(k)$ and $Y^N(k)$ respectively. We can then write,

$$\begin{aligned} E[|Z_x^N(k)|^2 | \mathcal{F}_k] &= \sum_{z_1=0}^{D^N(k)} \frac{\gamma z_1}{1 - \gamma B^N(k)} \left(1 - \frac{\gamma z_1}{1 - \gamma B^N(k)} \right) Y^N(k) \binom{D^N(k)}{z_1} \frac{1}{N} \left(1 - \frac{1}{N} \right)^{D^N(k) - z_1} \\ &+ \sum_{z_2=0}^{X^N(k) + Y^N(k)} \frac{\lambda \alpha z_2}{N^2} \left(1 - \frac{\lambda \alpha z_2}{N^2} \right) (A_k - X_k) \binom{X^N(k) + Y^N(k)}{z_2} \frac{1}{N} \left(1 - \frac{1}{N} \right)^{X^N(k) + Y^N(k) - z_2} \end{aligned}$$

$$\begin{aligned}
E[|Z_y^N(k)|^2 | \mathcal{F}_k] &= \sum_{z_1=0}^{D^N(k)} \frac{\gamma z_1}{1 - \gamma B^N(k)} \left(1 - \frac{\gamma z_1}{1 - \gamma B^N(k)}\right) Y^N(k) \binom{D^N(k)}{z_1} \frac{1}{N_1} \left(1 - \frac{1}{N}\right)^{D^N(k) - z_1} \\
&+ \sum_{z_2=0}^{X^N(k) + Y^N(k)} \frac{\lambda \sigma z_2}{N^2} \left(1 - \frac{\lambda \sigma z_2}{N^2}\right) (N - A_k - Y_k) \binom{X^N(k) + Y^N(k)}{z_2} \frac{1}{N_2} \left(1 - \frac{1}{N}\right)^{X^N(k) + Y^N(k) - z_2}
\end{aligned}$$

We see that the above two terms can be bound by constants c_x and c_y respectively and we can see that the noise conditions are satisfied.

(iv) **Convergence of initial conditions** By choice, we have $\tilde{X}^N(0) = x(0)$ and $\tilde{Y}^N(0) = y(0)$.

Thus by Theorem 7, we have for each $T > 0$ and each $\epsilon > 0$,

$$P\left(\sup_{0 \leq t \leq T} \left\| (\tilde{B}^N(\lfloor Nt \rfloor), \tilde{D}^N(\lfloor Nt \rfloor), \tilde{X}^N(\lfloor Nt \rfloor), \tilde{Y}^N(\lfloor Nt \rfloor)) - (b(t), d(t), x(t), y(t)) \right\| > \epsilon \right) \xrightarrow{N \rightarrow \infty} 0$$

where $(b(t), d(t), x(t), y(t))$ is the unique solution of the ODE

$$\begin{aligned}
\dot{b} &= d \\
\dot{d} &= \frac{\Gamma d}{1 - \Gamma b} (1 - b - d) - d \\
\dot{x} &= \lambda \alpha(x + y)(a - x) + \frac{\Gamma d}{1 - \Gamma b} y \\
\dot{y} &= \lambda \sigma(x + y)(1 - a - y) - \frac{\Gamma d}{1 - \Gamma b} y
\end{aligned}$$

with initial conditions $(b(0) = 0, d(0) = d_0, x(0) = x_0, y(0) = 0)$. ■

APPENDIX G SOLUTION OF THE ODE

Given the system of ODEs

$$\begin{aligned}
\dot{b} &= d \\
\dot{d} &= -d + \frac{\Gamma d}{1 - \Gamma b} (1 - b - d)
\end{aligned}$$

Dividing \dot{d} by \dot{b} ,

$$\frac{dd}{db} = \frac{\Gamma \dot{b} - \dot{b} - \Gamma \dot{b}^2}{d(1 - \Gamma b)}$$

By separating the variables,

$$\frac{dd}{\Gamma - 1 - \Gamma d} = \frac{db}{1 - \Gamma b}$$

Integrating on both sides, and after taking anti-logarithm

$$\Gamma - 1 - \Gamma d = c_1(1 - \Gamma b)$$

Differentiating on both sides yields $\dot{d} = c_1 d$ and hence $d(t) = c_2 e^{c_1 t}$. Substituting in the above equation for $d(t)$ we get

$$b(t) = \frac{c_2}{c_1} e^{c_1 t} + \frac{1 + c_1 - \Gamma}{\Gamma c_1}$$

Solving for constants using the initial conditions $b(0) = 0, d(0) = d_0$ results in $c_1 = -(1 + \Gamma d_0 - \Gamma) =: -r$ and $c_2 = d_0$. Hence we have,

$$\begin{aligned}
b(t) &= \frac{d_0}{r} - \frac{d_0}{r} e^{-rt} \\
d(t) &= d_0 e^{-rt}
\end{aligned}$$
■

APPENDIX H
CONVERGENCE OF $h_\gamma^{(N)}(k)$ TO b_∞

The solution of the ODE suggests that $\lim_{t \rightarrow \infty} b(t) = d_0/r$. This is consistent with the fact that $\lim_{t \rightarrow \infty} \frac{h_\gamma^{(N)}(k)}{N} \rightarrow \frac{d_0}{1-(1-d_0)\Gamma} = b_\infty$ as we now proceed to show.

$$h_\gamma^{(N)}(k) = k \left[1 + (N-k)\gamma \left[1 + (N-k-1)\gamma \left[1 + \dots \right. \right. \right.$$

$$\left. \left. \left. h_\gamma^{(N)}(k+1) = (k+1) \left[1 + (N-k-1)\gamma \left[1 + \dots \right. \right. \right. \right. \right.$$

Thus we can write,

$$h_\gamma^{(N)}(k) = k \left[1 + \gamma(N-k) \frac{h_\gamma^{(N)}(k+1)}{k+1} \right]$$

Now substituting for $k = d_0 N$ and noting that $\Gamma = \gamma N$ we have

$$\frac{h_\gamma^{(N)}(Nd_0)}{Nd_0} = 1 + \Gamma(1-d_0) \frac{h_\gamma^{(N)}(N(d_0 + \frac{1}{N}))}{N(d_0 + \frac{1}{N})}$$

Taking $N \rightarrow \infty$ and noting that $h_\gamma^{(N)}(k)$ is a continuous function, we have

$$\frac{1}{d_0} \frac{h_\gamma^{(N)}(Nd_0)}{N} = 1 + \frac{\Gamma(1-d_0)}{d_0} \frac{h_\gamma^{(N)}(Nd_0)}{N}$$

Take limits on both sides, and solving for the unknown,

$$\frac{h_\gamma^{(N)}(Nd_0)}{N} \rightarrow \frac{d_0}{1 - (1-d_0)\Gamma} = b_\infty$$

■